

DOCUMENT RESUME

ED 179 171

HE 012 002

AUTHOR McGee, Reece
TITLE Criteria Problems in Assessing Teaching Performance.
PUB DATE 79
NOTE 9p.: Paper presented at the annual meeting of the American Sociological Association (1979)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Accountability; College Environment; *College Faculty; Educational Responsibility; *Effective Teaching; *Evaluation Criteria; Higher Education; Models; *Researchers; *Teacher Evaluation; *Teacher Role; Values

ABSTRACT

The need to consider teaching evaluation criteria within the social structural and normative background within which college teaching takes place is discussed, along with the need to consider the substantive content of the criteria. In research-oriented American universities many feel that one cannot be promoted for teaching because institutional reward is positively correlated with research publication and other scholarly activity. One reason for this phenomena is that there is no tradition for concern with quality of instruction. A second structural phenomenon is that the major universities of the United States adopted the German model of scholarship, a disciplinary model, with the laboratory and speculative investigation at the center of the faculty member's role rather than the transmission of knowledge to students. The third phenomenon is that during the great research expansion of World War II, and the two decades following it, the academic profession and institutions of America adopted the researcher role as the academic role model. The research-oriented value system might be amenable to change if an appropriate methodology for the evaluation of teaching were invented. One of the general problems regarding the assessment of teaching is the absence of conventional agreement on what constitutes good teaching. A possible approach to this problem is to list behaviors essential for minimal satisfaction in teaching performance. (SW)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED179171

Criteria Problems in Assessing Teaching Performance

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Reece
McGee

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Reece McGee
Purdue University

Session on Issues in Teaching:
Evaluating and Rewarding Teaching

Annual Meeting 1979
American Sociological Association

HE 012 002

Criteria Problems in Assessing Teaching Performance

It is the contention of this paper that a principal source of problems in rewarding teaching within American higher education is the lack of criteria for evaluating it which are socially or politically, rather than technically, acceptable to academics. If this is the case, then any consideration of teaching evaluation criteria must consider the social structural and valuational or normative background within which college teaching takes place as well as the substantive content of the criteria under debate. I will try to suggest here what that background may be and then offer what I believe may be a tactical entering wedge for avoiding or displacing it. The paper's major address is to research-oriented universities because that is where the problem is most intense and most readily visible, as well as where it mainly originated, but the situation in other kinds of institution will be noted.

In the research-oriented universities -- that is, the major universities of the United States -- it is a common complaint that one "cannot be promoted for teaching," i.e., that institutional reward in the form of promotion, and perhaps to a lesser extent in pay, is positively correlated with research publication and other "scholarly" activity and that teaching goes relatively unrewarded. In its extreme form the complaint may even assert that teaching correlation between them and reward. The perceived problem in institutions of this type, then, is the actual activities and consequences of institutionalization of the "publish or perish" metaphor.

In other types of institutions, while there may be no perception that faculty members are subjected to negative sanctions for taking their teaching tasks seriously, it is complained that teaching evaluation is unsystematic, that criteria for assessment are vague, that too much or too little attention is paid to classroom student evaluation, that promotion committees rely on hearsay, and so forth. The ASA Projects on Undergraduate Teaching have documented such allegations for institutions of all types.¹

From what I know of the academic profession and its literature, there may be at least four social structural phenomena that contribute to an explanation of these conditions, each with normative or valuational baggage that makes it difficult to alter their effects. Each has differential salience for different types of institution and the normative structures they produce appear in different forms among the different types. Nonetheless the four form a cultural matrix in American higher education which, in its various manifestations, seems to me to pose the problem to be addressed here.

The first of these phenomena is simply tradition. There is no tradition in American universities for concern with quality of instruction. Rather it has been taken for granted that an instructor's mastery of subject matter was all that was required to assure adequacy in teaching. This may stem in part from our profession's early adoption of the British collegial model of faculty governance (in the absence of the actual "community of scholars" represented by the English college). It may also be in part the result of the fact that, for so many of the early years of higher education in the United States, college and, later, university, instruction were seen essentially as a "finishing" process, the object of which was the production of "well-rounded"

gentlemen on the British upper-class model. The purpose of education, in other words, was exposure to high culture, which was then presumably absorbed through note learning and intellectual osmosis.

A second structural phenomenon is that the major universities of the United States adopted the German model of scholarship as they came into being, a disciplinary model, with the laboratory and speculative investigation at the center of the faculty member's role rather than the transmission of knowledge to (primarily) undergraduate students. Our universities thus adopted a value system in being which defined appropriate role models for teachers to evaluate and research scholarship as the only activity truly appropriate for the professoriat. It is a model which effectively intrudes professional and disciplinary values (i.e., extra-institutional norms) into the organization we call a university. Evidence in support of this explanation seems to lie in the fact that the disciplines of the physical sciences were those first imported from Germany, and it is still among those fields where the "research uber alles" value system remains most strongly pronounced today.

The third phenomenon is that during the great research expansion of World War II and the two decades following it, the academic profession and institutions of America adopted as normal a set of beliefs, values, and practices adapted to an historical context which was, in fact, distinctly abnormal. Only the very senior and the very junior among us today have experienced as typical in their own professional lives what the rest of us are prone to call the contemporary "recession" in academics. But the time of financial plenty -- a brief, necessarily temporary, and perhaps somewhat chimerical Golden Age -- in fact was never typical of educational conditions in this country. The present condition of scarce jobs and resources for higher education is historically far more normal than the period between 1942 and 1971.

The effects of this period on teaching were disastrous. They helped produce a set of norms and valuational practices that made the quality of a faculty member's instruction irrelevant to his or her utility to the department and institution and damaged institutional capacity to operate. It also removed, perhaps forever, a significant degree of faculty control over institutional operation. What happened was simple: With the bounties of federal and foundation financing overflowing, the research entrepreneur became the academic role model. Only the researcher could tap the wellsprings of "outside money" to bring into the department and institution the luxuries of graduate appointments, extra secretaries, travel money, expensive hard and soft ware, and so forth, impossible to support from more limited appropriations or private endowment. At the same time the vast enrollment waves first of returning veterans and then the children of the Baby Boom created added tuition resources and ensured that our classrooms would be filled no matter what we did, or failed to do, there. (This was essential because the ostensible public mission of the university remained that of instruction and the legislatures and donors continued to believe that that was what we were doing.)

The consequence of all this was that the value of the academic man or

woman became that of "reputation," national or, even and most desirable, international. Professional reputation was what brought money in, insured that ever more graduate students could be attracted, provided "overhead," etc. Forgotten in the rush to enhance departmental and institutional reputation through the recruitment of scholars of eminence were any considerations of quality of instruction, how the institutions would support themselves when the money and enrollments dried up, and the desiderata that a faculty be self-governing or administering. (There was no national reputation to be gained by individuals or departments from dedicated classroom instruction or committee work for the college. These were invisible and, hence, valueless activities, best left to administrators, teaching assistants, and untenured junior faculty members who were not, after all, yet of any prestige value to the department.)

This explanation, of course, best fits the major universities. The phenomenon occurred to some degree in institutions of other types as well, however, and was exacerbated in some of them by a fourth, harder to define or even label and difficult to demonstrate empirically. This is what could be called the secondary school orientation to education. A large number of contemporary colleges and universities began as Normal Schools, two-year institutions exclusively devoted to the training of elementary and secondary school teachers. Many other institutions had this function as one major role. Now due to a variety of historical phenomena, not the least of which is the governance of lower education by lay boards which very closely oversee it, elementary and secondary school teachers have never had the prestige or experienced the degree of professionalization of college teachers. One result is that they are treated by their institutions and executives much more as replaceable parts: One is as good as another, certification held constant. Their pay is lower, their tenure less secure and meaningful, their teaching activities more routinized, and their opportunities for originality, creativity, and control over their conditions of work much more restricted. They are hired and fired with monotonous regularity.

Most important for our purposes, in the collegiate institutions dominated by the preparation of students to become lower education faculty, the attitudes and values about teaching and teachers just described inevitably tainted appraisal and understanding of their own faculties and institutions. It is no accident that schools dominated by teacher training missions have traditionally had much more centralized administrative structures, more authoritarian governing practice, less academic freedom, and more capricious reward systems. To a degree they have modeled the systems they serve, with the result that the evaluation and reward of instruction there, too, was likely to suffer, although to be sure from somewhat different problems than those described by our third phenomenon. (It is my impression that to the extent that institutions of this type adopted the research model, they merely complicated the problem of assessing teacher performance without either improving or systematizing it. It would also be difficult to defend the proposition that they contributed substantially to the production of important research.)

It is my thesis that as a consequence of all of these phenomena, American higher education now collectively possesses a nationwide, conventionalized, readily-accessible model for the evaluation of academic performance, a model on which there is a considerable degree of consensus. It varies somewhat by discipline and institutional type, but its overall effect is to denigrate teaching by ignoring or trivializing it. It also has the effect of making the

curriculum vitae into the conventionalized mechanism through which rewards are systematically allocated. The academic value system, in abort, created the material method and means for its own actualization in reward-granting behavior. In less research-oriented institutions the content of the paper record may be less substantially devoted to research accomplishment, but actual quality in teaching performance is unlikely to be measured because that is not the variable at issue. Compliance and regularity are more likely to be rewarded in the lower education model, although phrased in the record as devotion to teaching. (And it may be worth noting in passing that quality of research cannot be demonstrated by the paper record in the major institutions either, quantity being the variable normally rewarded.)

If the thesis just suggested is accepted, two hypotheses seem logically to follow. The first is that the existence of the scholarship model, and the institutionalized value system which supports it, inhibits the evolution of other models. It works too well; it is familiar, conventional; we know how to use it. And teaching activity does not fit; the scholarship model is inadequate to measure or demonstrate quality in teaching endeavor and the value system does not support it as important in any event.

The second hypothesis is that under the present circumstances of increases in extra-institutional pressures for accountability that urge the universities to respond more directly to student clientele, the research-oriented value system might be amendable to change if an appropriate methodology for the evaluation of teaching were invented, if the curriculum vitae could be replaced or supplemented by other mechanisms of assessment. It may be, in other words, that in part the universities do not now reward teaching more than they do because they do not know how to do so. The only reward-allocation mechanism they possess is unsuited to it, having been invented for the evaluation of other activities. But if an appropriate one were to be made available, its existence might result in its utilization.

This suggestion, however, raises a most difficult issue. There is probably no procedure or paper record that can measure or demonstrate teaching quality as that term is normally used. The reason, of course, is that we have no conventional agreement on what constitutes "good teaching." A procedure could be invented to produce a curriculum vitae in which teaching and teaching-related activities could be described in somewhat the same way as research activities and publications, and in which various estimations of teaching quality could appear. Such a document could demonstrate that an individual put serious effort into teaching, communicated his or her interests to other colleagues, had imaginative classroom presentations, was liked by students who thought well of their experience in his or her classes, and was approved as effective and enthusiastic by peers. But it would be flawed by the absolute lack of agreement on what constitutes "good teaching" and the difficulty of demonstrating in conventionalized operations such qualities as may be thought to make it up.

Thus one of the major elements contributing to the general problem of criteria for assessment of teaching is that in the absence of anything even approaching conventional agreement on what constitutes "good teaching" in a given field, it is impossible to reward people for doing it. One reason

for this is the relative invisibility of the product of instruction. Teaching activity is as readily observable as research activity, if not more so, because it is more public, being performed directly before an audience. But instructional results are not. (Although there are ways of making them so; see Wilson and Goldsmid, 1979.) Research results in some kind of material product: A paper which can be judged by a national or even international audience, a book, an actual artifact, product or process that others can evaluate and try out for themselves. The results of an instructional experience, to the contrary, are most of them vague, subjective, and perhaps very long in becoming apparent.

There are, thus, two profound problems to be overcome if we wish to introduce the possibility of systematic reward for teaching activity; The problem of definition and the problem of consensus. It seems to me hopeless to seek agreement at this time on what constitutes "good" or "effective" teaching; neither Morse code nor Esperanto would be effective in the Tower of Babel. We might, however, be able to begin to define, and expect reasonably wide agreement on, some list of behaviors essential for minimal satisfaction in teaching performance. Such a list might be phrased in terms of observable behavior in classroom and office, and of student reactions to it. It is not unreasonable to propose, for example, that it might be considered essential to minimal satisfaction in instruction for a faculty member to announce and keep regularly a reasonable number of office hours per week. Or that some significant proportion of his or her students should agree that he or she usually appears to have prepared for classes beforehand. If such a list were relatively brief, operationalized where necessary, and eminently reasonable, it is not unrealistic to hope that it might achieve some effective consensus. Reasonability seems a critical quality. It would be impossible, of course, for an opponent to argue in favor of bad teaching, but the possibilities for quibbling are so great, and the opportunities so endless, that for maximum effectiveness any such device would have to be almost completely unexceptionable, at least upon first introduction. The key epistemological issue, of course, is that it would turn our attention away from the fool's errand of seeking a consensual definition of "good teaching" to the more practical and agreeable objective of defining minimal behavioral competence.

Lists or guidelines of the general type already exists.² And it would not be difficult for a department to create its own. Despite the widespread belief that there is no consensus in our own field, at least, and hence that any agreement on such minimum essentials would be impossible, a variety of evidence exists to the contrary. The immense uniformity of our introductory textbooks, for example, and our acceptance and use of them in astronomical numbers, indicates a wide consensus on what that course should cover. Similar is the University of North Carolina department's list of minimal essential learnings for the introductory course.³ I have little doubt that it would be relatively easy to produce similar lists for instructor behavior which could achieve wide acceptance. It appears to me that some such means are the way around the roadblocks now standing in the way of developing criteria for assessing teacher performance. Certainly we will gain nothing by continuing to wring our hands over the "impossibility" of doing so. The social/political obstacles created by our value system are quite real, but are historical accidents, products of times and circumstances gone and never to return. The major mission of all of our educational institutions remain what it has always

been: The instruction of undergraduates. If we suffer from cognitive and other dissonance because our values do not reflect our actual activities and requirements, we have but to change them. The rest will fall into place by itself.

NOTES

1. Bowker, Lee (ed.)
1978
Data Book on the Institutional Context of Teaching Undergraduate Sociology. ASA Project on Teaching Undergraduate Sociology. Washington, D.C. ASA Teaching Resources Center.
2. Helling, Barbara A.
1976
"Looking for Good Teaching: A Guide to Peer Observation." Excerpt from a Danforth Faculty Fellowship Report. Distributed by the ASA Teaching Resources Center, Washington, D.C.
3. Wilson, E.K. and C.A. Goldsmid
1979
Sociology in Passage, Belmont, CA: Wadsworth. 1980